Tight Results on Multiregister Fourier Sampling: Quantum Measurements for Graph Isomorphism Require Entanglement

Cristopher Moore
moore@cs.unm.edu
Department of Computer Science
University of New Mexico

Alexander Russell
acr@cse.uconn.edu
Department of Computer Science and Engineering
University of Connecticut

February 1, 2008

Abstract

We establish a general method for proving bounds on the information that can be extracted via arbitrary entangled measurements on tensor products of hidden subgroup coset states. When applied to the symmetric group, the method yields an $\Omega(n \log n)$ lower bound on the number of coset states over which we must perform an entangled measurement in order to obtain non-negligible information about a hidden involution. These results are tight to within a multiplicative constant and apply, in particular, to the case relevant for the Graph Isomorphism problem.

Part of our proof was obtained after learning from Hallgren, Rötteler, and Sen that they had obtained similar results.

1 Introduction: the hidden subgroup problem

Many problems of interest in quantum computing can be reduced to an instance of the *Hidden Subgroup Problem* (HSP). This is the problem of determining a subgroup H of a group G given oracle access to a function $f : G \to S$ with the property that $f(g) = f(hg) \Leftrightarrow h \in H$. Equivalently, f is constant on the cosets of H and takes distinct values on distinct cosets.

All known efficient solutions to the problem rely on the *standard method* [4], in which we prepare a uniform superposition over the elements of G and measure the value of the oracle on this superposition. This yields a uniform superposition over a uniformly random left coset, $|cH\rangle = (1/\sqrt{|H|}) \sum_{h \in H} |ch\rangle$, or equivalently a mixed state, $\rho_H = (1/|G|) \sum_{c \in G} |cH\rangle \langle cH|$. The question is how much information about the subgroup H can be gained by measuring this state. *Fourier sampling* measures ρ_H according to the Fourier basis, i.e., according to the irreducible representations of G; as we discuss below, the optimal measurement is always of this type.

History of the Hidden Subgroup Problem. Both Simon's and Shor's seminal algorithms rely on the standard method over an abelian group. In Simon's problem [36], $G = \mathbb{Z}_2^n$ and f is an oracle such that, for some y, $f(x) = f(x + y)$ for all x; in this case $H = \{0, y\}$ and we wish to identify y. In Shor's factoring algorithm [35] G is (essentially) the group $\mathbb{Z}_n^*$ where n is the number we wish to factor, $f(x) = r^x \bmod n$ for a random $r < n$, and H is the subgroup of $\mathbb{Z}_n^*$ whose index is the multiplicative order of r.

While the *nonabelian* hidden subgroup problem appears to be much more difficult, it has very attractive applications. In particular, solving the HSP for the symmetric group S_n would provide an efficient quantum algorithm for the Graph Automorphism and Graph Isomorphism problems (see e.g. Jozsa [19] for a review). Another important motivation is the relationship between the HSP over the dihedral group with hidden shift problems [6] and cryptographically important cases of the Shortest Lattice Vector problem [32].

So far, algorithms for the HSP are only known for a few families of nonabelian groups [33, 18, 9, 25, 16, 2]. Ettinger and Høyer [7] provided another type of result (see also [31]) by showing that Fourier sampling can solve the HSP for the dihedral groups D_n in an *information-theoretic* sense. That is, a polynomial number of experiments gives enough information to reconstruct the subgroup, though it is unfortunately unknown how to determine H from this information in polynomial time.

To discuss Fourier sampling for a nonabelian group G, one needs to consider *representations* of the group, namely homomorphisms $\rho : G \to \mathsf{U}(V)$ where $\mathsf{U}(V)$ is the group of unitary matrices acting on some $\mathbb{C}$-vector space V of dimension d_ρ. It suffices to consider *irreducible* representations, namely those for which no nontrivial subspace of V is fixed by the various operators $\rho(g)$. Once a basis for each irreducible ρ is chosen, the matrix elements ρ_{ij} provide an orthogonal basis for the vector space of all $\mathbb{C}$-valued functions on G. The quantum Fourier transform then consists of transforming (unit-length) vectors in $\mathbb{C}[G] = \{\sum_{g \in G} \alpha_g |g\rangle \mid \alpha_g \in \mathbb{C}\}$ from the basis $\{|g\rangle \mid g \in G\}$ to the basis $\{|\rho, i, j\rangle\}$ where ρ is the name of an irreducible representation and $1 \le i, j \le d_\rho$ index a row and column (in a chosen basis for V). Indeed, this transformation can be carried out efficiently for a wide variety of groups [3, 15, 24].

A basic question concerning the hidden subgroup problem is whether there is always a basis for the representations of G such that measuring in this basis provides enough information to determine the subgroup H. This framework is known as *strong Fourier sampling*. In [28], Moore, Russell and Schulman answered this question in the negative, showing that subgroups of S_n relevant to Graph Isomorphism cannot be determined by this process; more generally, they showed that no subexponential number of positive operator-valued measurements (POVMs) of individual coset states suffices.

The next logical step is to consider *multi-register* algorithms, in which we prepare multiple coset states and subject them to *entangled* measurements. Ettinger, Høyer and Knill [8] showed that the HSP on arbitrary groups can be solved information-theoretically with a polynomial number of registers, and the authors of this article have shown how to carry out such a measurement for the case relevant to Graph Isomorphism in the Fourier basis [26]. For the dihedral group D_n, Ip [17] showed that the optimal measurement for two registers is entangled, and Kuperberg [23] devised a subexponential ($2^{O(\sqrt{\log n})}$) algorithm that works by performing entangled measurements on two registers at a time. Bacon, Childs, and van Dam [1, 2] determined the optimal multiregister measurement for certain metabelian groups, and use this to devise the first efficient multiregister algorithms. The present authors have generalized these optimality results to the case where H and G form a Gel'fand pair [27].

Our contribution. Whether a similar approach can be applied to the symmetric group, offering an efficient algorithm for Graph Isomorphism, is the principal open question in this area. Here we establish a general method for bounding the information that can be extracted by arbitrary entangled measurements on tensor products of coset states. These bounds give rise to the following theorem:

Theorem 1. *Suppose we are given the coset state $\rho_H^{\otimes k}$ on k registers for the hidden subgroup $H = \{1, m\}$ where m is chosen uniformly at random from a conjugacy class M of involutions. Given that we observe the representation $\boldsymbol{\rho} = \rho_1 \otimes \cdots \otimes \rho_k$, let B be a basis for $\boldsymbol{\rho}$, let $\mathcal{H}_m(\mathbf{b})$ be the probability that we observe the vector $\mathbf{b} \in B$, and let $\mathcal{U}$ be the uniform distribution on B. Then there is a constant $C > 0$ such that, if $k < Cn \log_2 n$, with probability $1 - n^{-\Omega(n)}$ in m and $\boldsymbol{\rho}$, we have*

$$\|\mathcal{H}_m - \mathcal{U}\|_1 = n^{-\Omega(n)} \ .$$

Thus, unless $k = \Omega(n \log n)$, it takes a superpolynomial number of experiments to distinguish the different subgroups $H = \{1, m\}$ from each other, or from the trivial subgroup, for which the observed distribution is uniform. Along with the fact that $O(n \log n)$ registers suffice [8, 26], this shows that entangled measurements over $\Theta(n \log n)$ registers are both necessary and sufficient.

Note that this result is much stronger than the claim that the total query complexity of this case of the Hidden Subgroup Problem is $\Theta(n \log n)$ (where each query consists of generating a coset state); indeed, one can immediately obtain $\Omega(n)$ lower bounds on the query complexity of determining an involution m by embedding $\mathbb{Z}_2^n$ into S_{2n}. In fact, these bounds can be obtained even without the assumption that each query generates a coset state [22]. The query complexity of the decision problem of whether H is of the form $\{1, m\}$ or is trivial was recently shown to be $\Omega(n)$ in a natural hidden shift model [5].

Such query complexity lower bounds, however, do not preclude the possibility of using multiple independent applications of (single-register) Fourier sampling to solve the problem; for instance, in the dihedral group, each such measurement yields a constant amount of information [7]. In contrast, the result proved here shows that in order to gain non-negligible information about the hidden subgroup, and thus about whether the two graphs are isomorphic or not, one must measure $O(\log |G|)$ registers simultaneously in an entangled basis. This greatly restricts the set of possible quantum algorithms for Graph Isomorphism.

Remark. A preliminary version of this paper appeared in [29] where we developed a general framework for bounding the available information in the multiregister case, including Lemmas 2–5, and showed that entangled measurements over two registers are insufficient. The proof of Lemma 9 below, on which Theorem 1 depends, was obtained after learning from Hallgren, Rötteler, and Sen that they had obtained results similar to Theorem 1 by building on the machinery of [29].

2 The structure of the optimal measurement

We focus on the special case of the hidden subgroup problem called the *hidden conjugate problem* in [25]. Here there is a (non-normal) subgroup H, and we are promised that the hidden subgroup is one of its conjugates, $H^g = g^{-1}Hg$ for some $g \in G$; the goal is to determine which.

The most general possible measurement in quantum mechanics is a positive operator-valued measurement (POVM). It is easy to see [28] that the optimal POVM for the Hidden Subgroup Problem on a single coset state consists of measuring the name ρ of the irreducible representation, followed by a POVM on the vector space V on which ρ acts. For simplicity, here we will restrict ourselves to von Neumann measurements, in which we measure the space on which ρ acts according to some orthonormal basis B. As in [28, 29] our results can easily be extended to arbitrary POVMs.

Under Fourier sampling, the probability we observe ρ, and the conditional probability that we observe a given $\mathbf{b} \in B$, are given by

$$\mathcal{H}(\rho) = \frac{d_\rho |H|}{|G|} \mathbf{rk}\ \Pi_H \qquad \text{and} \qquad \mathcal{H}(\rho, \mathbf{b}) = \frac{\|\Pi_H \mathbf{b}\|^2}{\mathbf{rk}\ \Pi_H} \tag{2.1}$$

where Π_H is the projection operator $1/|H| \sum_{h \in H} \rho(h)$. When H is nontrivial, the probability distribution over B changes for a conjugate H^g in the following way:

$$\mathcal{H}(\rho, \mathbf{b}) = \frac{\|\Pi_H g\mathbf{b}\|^2}{\mathbf{rk}\ \Pi_H}$$

where we write $g\mathbf{b}$ for $\rho(g)\mathbf{b}$. In contrast, if H is the trivial subgroup, $\Pi_H = \mathbb{1}_{d_\rho}$ and $\mathcal{H}(\rho)$ is the *Plancherel distribution* $\mathcal{P}(\rho) = d_\rho^2/|G|$, and $\mathcal{H}(\rho, \mathbf{b}_j) = 1/d_\rho$ is the uniform distribution.

3 The expectation and variance of an involution projector

Definition 1. *Let ρ be a representation of a group G acting on a space V and let σ be an irreducible representation of G. We let $\mathfrak{I}_\sigma^\rho$ denote the projection operator onto the* σ-isotypic *subspace of V, the subspace spanned by all copies of σ in ρ. We remark that this projection operator can written as the sum $\mathfrak{I}_\sigma^\rho \mathbf{v} = \frac{d_\sigma}{|G|} \sum_g \chi_\sigma^*(g) g\mathbf{v}$, regardless of the structure of ρ. See, e.g., [34].*

The following two lemmas are proved in [28]; we repeat them here for convenience.

Lemma 2. *Let ρ be a representation of a group G acting on a space V and let $\mathbf{b} \in V$. Let m be chosen uniformly from a conjugacy class M of involutions. If ρ is irreducible, then*

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \frac{\chi_\rho(M)}{d_\rho} \|\mathbf{b}\|^2 \quad .$$

If ρ is reducible, then

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \sum_{\sigma \in \widehat{G}} \frac{\chi_\sigma(M)}{d_\sigma} \|\mathfrak{I}_\sigma^\rho \mathbf{b}\|^2 \quad .$$

Lemma 3. *Let ρ be a representation of a group G acting on a space V and let $\mathbf{b} \in V$. Let m be chosen uniformly from a conjugacy class M of involutions. Then*

$$\mathrm{Exp}_m |\langle \mathbf{b}, m\mathbf{b} \rangle|^2 = \sum_{\sigma \in \widehat{G}} \frac{\chi_\sigma(M)}{d_\sigma} \left\| \mathfrak{I}_\sigma^{\rho \otimes \rho^*} (\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \quad .$$

Now, given an involution m and the hidden subgroup $H = \{1, m\}$, let $\Pi_m = \Pi_H$ denote the projection operator given by $\Pi_m \mathbf{v} = (1/2)(\mathbf{v} + m\mathbf{v})$. Then the expectation and variance of $\|\Pi_m \mathbf{b}\|^2$ are given by the following lemma, also from [28].

Lemma 4. *Let ρ be an irreducible representation acting on a space V and let $\mathbf{b} \in V$. Let m be chosen uniformly from a conjugacy class M of involutions. Then*

$$\mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2 = \frac{1}{2}\|\mathbf{b}\|^2 \left(1 + \frac{\chi_\rho(M)}{d_\rho}\right) \tag{3.1}$$

$$\mathrm{Var}_m \|\Pi_m \mathbf{b}\|^2 \leq \frac{1}{4}\sum_{\sigma \in \widehat{G}} \frac{\chi_\sigma(M)}{d_\sigma} \left\| \mathfrak{J}_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \quad . \tag{3.2}$$

Finally, we point out that since $\mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2 = \|\mathbf{b}\|^2 \frac{\mathbf{rk}\,\Pi_m}{d_\rho}$ we have

$$\frac{\mathbf{rk}\,\Pi_m}{d_\rho} = \frac{1}{2}\left(1 + \frac{\chi_\rho(M)}{d_\rho}\right) \quad . \tag{3.3}$$

4 Variance and decomposition for multiregister experiments

We turn now to the multi-register case, where Steps 1, 2 and 3 are carried out on k independent registers. This yields a state in $\mathbb{C}[G^k]$, i.e., $|c_1 H\rangle \otimes \cdots \otimes |c_k H\rangle$ where the c_i are uniformly random coset representatives. The symmetry argument of [28] applies to each register, so that the optimal measurement is consistent with first measuring the representation name in each register. However, the optimal measurement generally does not consist of k independent measurements on this tensor product state; rather, it is entangled, consisting of measurement in a basis whose basis vectors $\mathbf{b}$ are not of the form $\mathbf{b}_1 \otimes \cdots \otimes \mathbf{b}_k$.

In this section, we extend the results of [28] to the case of multiple coset states in three steps. First, in Section 4.1, we generalize the expressions of Lemma 4 for the expectation and variance of the observed distribution to the multiregister case. In Sections 4.2 and 4.3, we bound the expectation and variance of the probability distribution, by controlling to what extent tensor product vectors project into "bad" low-dimensional representations with large normalized characters. These bounds are far tighter than those in [28, 29], in which we pessimistically bounded these projections simply by estimating the multiplicity of bad representations. Finally, in Section 4.4, we combine these bounds to bound the expectation over $\boldsymbol{\rho}$ of the total variation distance between the observed distribution and the uniform distribution.

4.1 Variance for Fourier sampling product states

We begin by generalizing Lemmas 1, 2, and 3 of [28] to the multi-register case. The reasoning is analogous to that of Section 4 of [28]; the principal difficulty is notational, and we ask the reader to bear with us.

We assume we have measured the representation name on each of the registers, and that we are currently in an irreducible representation of G^k labeled by $\boldsymbol{\rho} = \rho_1 \otimes \cdots \otimes \rho_k$. For a subset $S \subset [k]$, let us introduce the shorthand $\rho_S = \otimes_{i \in S} \rho_i$ and $\rho_S \otimes \mathbb{1} = \bigotimes_{i \in S} \rho_i \otimes \bigotimes_{i \in \overline{S}} \mathbb{1}$, operating in the natural way on the vector space that supports $\boldsymbol{\rho}$.

Then given a subset $I \subseteq [k]$, we can separate this tensor product into the registers inside I and those outside, and then decompose the product of those inside I into irreducibles:

$$\boldsymbol{\rho} = \rho_I \otimes \rho_{\overline{I}} = \left(\bigoplus_{\sigma \in \widehat{G}} a_\sigma^I \sigma\right) \otimes \rho_{\overline{I}}$$

where a_σ^I is the multiplicity of σ in ρ_I. Now given an irreducible representation σ, we let $\Pi_\sigma^I = \mathfrak{I}_\sigma^{\rho_I \otimes \mathbb{1}}$ denote the projection operator onto the subspace acted on by $a_\sigma^I \sigma \otimes \rho_{\overline{I}}$. In other words, Π_σ^I projects the registers in I onto the subspaces isomorphic to σ, and leaves the registers outside I untouched. Note that in the case where I is a singleton we have $\mathfrak{I}_{\rho_i}^{\rho_i \otimes \mathbb{1}} = \mathbb{1}$.

As before, the hidden subgroup is $H = \{1, m\}$ for an involution m chosen at random from a conjugacy class M. However, we now have, in effect, the subgroup $H^k \subset G^k$, and summing over the elements of H^k gives the projection operator $\Pi_{H^k} = \Pi_m^{\otimes k}$. The probability of observing a representation $\boldsymbol{\rho}$ under weak sampling is thus

$$\mathcal{H}(\boldsymbol{\rho}) = \mathcal{H}_M^{\otimes k}(\boldsymbol{\rho}) \triangleq \frac{d_{\boldsymbol{\rho}} |H|^k}{|G|^k} \left(\mathbf{rk}\ \Pi_H\right)^k \ .$$

Conditioned upon observing $\boldsymbol{\rho}$, the probability we observe an (arbitrarily entangled) basis vector $\mathbf{b} \in \boldsymbol{\rho}$ is

$$\mathcal{H}(\boldsymbol{\rho}, \mathbf{b}) = \mathcal{H}_m^{\otimes k}(\boldsymbol{\rho}, \mathbf{b}) \triangleq \frac{\left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2}{\mathbf{rk}\ \Pi_m^{\otimes k}} \ . \tag{4.1}$$

As indicated, we elide the superscripts and subscripts when they can be inferred from context. We remark that the distribution $\mathcal{H}^{\otimes k}(\boldsymbol{\rho})$ depends only on M and can be written as a product distribution: $\mathcal{H}^{\otimes k}(\boldsymbol{\rho}) = \prod_i \mathcal{H}^{\otimes 1}(\rho_i)$. The distribution $\mathcal{H}_m^{\otimes k}(\boldsymbol{\rho}, \mathbf{b})$, on the other hand, cannot in general be decomposed in this way as we consider arbitrarily entangled bases of $\boldsymbol{\rho}$ as opposed to product bases.

When we calculate the expectation of this over m, we will find ourselves summing the following quantity over the subsets $I \subseteq [k]$:

$$E^I(\mathbf{b}) \triangleq \sum_{\sigma \in \widehat{G}} \frac{\chi_\sigma(M)}{d_\sigma} \left\| \Pi_\sigma^I \mathbf{b} \right\|^2 \tag{4.2}$$

with $E^\emptyset(\mathbf{b}) = \|\mathbf{b}\|^2$ (since an empty tensor product gives the trivial representation). Note that $E^I(\mathbf{b})$ is real, since $\chi_\sigma(m)$ is real for any involution m.

For the variance, we will consider pairs of subsets $I_1, I_2 \subseteq [k]$ and decompositions of the form

$$\boldsymbol{\rho} \otimes \boldsymbol{\rho}^* = \left(\rho_{I_1} \otimes \rho_{I_2}^*\right) \otimes \left(\rho_{\overline{I}_1} \otimes \rho_{\overline{I}_2}^*\right) = \left(\bigoplus_{\sigma \in \widehat{G}} a_\sigma^{I_1, I_2} \sigma \right) \otimes \left(\rho_{\overline{I}_1} \otimes \rho_{\overline{I}_2}^*\right)$$

just as we considered $\rho \otimes \rho^*$ in the one-register case [28]. We then define the projection operator $\Pi_\sigma^{I_1, I_2} = \mathfrak{I}_\sigma^{(\rho_{I_1} \otimes \mathbb{1}) \otimes (\rho_{I_2} \otimes \mathbb{1})^*}$ onto the subspace acted on by $a_\sigma^{I_1, I_2} \sigma \otimes \rho_{\overline{I}_1} \otimes \rho_{\overline{I}_2}^*$ and we define the following quantity,

$$E^{I_1, I_2}(\mathbf{b}) \triangleq \sum_{\sigma \in \widehat{G}} \frac{\chi_\sigma(M)}{d_\sigma} \left\| \Pi_\sigma^{I_1, I_2} (\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \tag{4.3}$$

with $E^{\emptyset, \emptyset}(\mathbf{b}) = \|\mathbf{b}\|^4$.

We can now state the following lemma: note that (4.5) corresponds to (3.2) in the one-register case.

Lemma 5. *Let $\mathbf{b} \in \boldsymbol{\rho}$ and let m be chosen uniformly from a conjugacy class M of involutions. Then*

$$\mathrm{Exp}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2 = \frac{1}{2^k} \left(1 + \sum_{I \subseteq [k] : I \neq \emptyset} E^I(\mathbf{b}) \right) \ , \tag{4.4}$$

$$\mathrm{Var}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2 \leq \frac{1}{4^k} \sum_{I_1, I_2 \subseteq [k] : I_1, I_2 \neq \emptyset} E^{I_1, I_2}(\mathbf{b}) \ . \tag{4.5}$$

Proof. Let m^I denote the operator that operates on the ith register by m for each $i \in I$ and leaves the other registers unchanged. This acts on $\mathbf{b}$ as $\rho_I(m)$, and Lemma 2 implies that $\mathrm{Exp}_m\langle \mathbf{b}, m^I\mathbf{b}\rangle = E^I(\mathbf{b})$. Then (4.4) follows from the observation that

$$\Pi_m^{\otimes k}\mathbf{b} = \frac{1}{2^k}\sum_{I\subseteq[k]} m^I\mathbf{b}$$

and so

$$\mathrm{Exp}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2 = \mathrm{Exp}_m\langle \mathbf{b}, \Pi_m^{\otimes k}\mathbf{b}\rangle = \frac{1}{2^k}\sum_{I\subseteq[k]} \mathrm{Exp}_m\langle \mathbf{b}, m^I\mathbf{b}\rangle = \frac{1}{2^k}\sum_{I\subseteq[k]} E^I(\mathbf{b}) \ .$$

Separating out the term $E^{\emptyset}(\mathbf{b}) = \|\mathbf{b}\|^2$ completes the proof of (4.4).

Similarly, let the operator m^{I_1,I_2} act on $\mathbf{b}\otimes\mathbf{b}^*$ by multiplying the ith register of $\mathbf{b}$ by m whenever $i \in I_1$, multiplying the ith register of $\mathbf{b}^*$ whenever $i \in I_2$, and leaving the other registers of $\mathbf{b}$ and $\mathbf{b}^*$ unchanged. Then it acts as $(\rho_{I_1}\otimes\rho^*_{I_2})(m)$, and Lemma 2 implies $\mathrm{Exp}_m\langle \mathbf{b}\otimes\mathbf{b}^*, m^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\rangle = E^{I_1,I_2}(\mathbf{b})$. Then analogous to Lemmas 3 and 4, the second moment is

$$\begin{aligned}\mathrm{Exp}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^4 &= \mathrm{Exp}_m\langle \mathbf{b}, \Pi_m^{\otimes k}\mathbf{b}\rangle\langle \mathbf{b}^*, \Pi_m^{\otimes k}\mathbf{b}^*\rangle = \mathrm{Exp}_m\langle \mathbf{b}\otimes\mathbf{b}^*, (\Pi_m^{\otimes k}\otimes\Pi_m^{\otimes k})(\mathbf{b}\otimes\mathbf{b}^*)\rangle \\ &= \frac{1}{4^k}\sum_{I_1,I_2\subseteq[k]} \mathrm{Exp}_m\langle \mathbf{b}\otimes\mathbf{b}^*, m^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\rangle = \frac{1}{4^k}\sum_{I_1,I_2\subseteq[k]} E^{I_1,I_2}(\mathbf{b})\end{aligned}$$

and so the variance is

$$\begin{aligned}\mathrm{Var}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2 &= \mathrm{Exp}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^4 - \left(\mathrm{Exp}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2\right)^2 \\ &= \frac{1}{4^k}\sum_{I_1,I_2\subseteq[k]} \big(E^{I_1,I_2}(\mathbf{b}) - E^{I_1}(\mathbf{b})E^{I_2}(\mathbf{b})\big) = \frac{1}{4^k}\sum_{I_1,I_2\neq\emptyset} E^{I_1,I_2}(\mathbf{b}) - \frac{1}{4^k}\left|\sum_{I\neq\emptyset} E^I(\mathbf{b})\right|^2 \qquad (4.6)\end{aligned}$$

where we use the fact that $E^{I,\emptyset}(\mathbf{b}) = E^{\emptyset,I}(\mathbf{b}) = E^I(\mathbf{b})\,\|\mathbf{b}\|^2 = E^I(\mathbf{b})$. Finally (4.5) follows by neglecting the negative term of (4.6). □

As in the case of (one-register) Fourier sampling [28], the Plancherel distribution $\mathcal{P}^{\otimes k}(\boldsymbol{\rho}) = d_{\boldsymbol{\rho}}/|G|^k$ will play a special role in the analysis. Note that $\mathcal{P}^{\otimes k}(\boldsymbol{\rho}) = \prod \mathcal{P}(\rho_i)$ and that, consistent with our conventions for $\mathcal{H}$, we elide the superscript when it will cause no confusion.

In the following two sections, we establish bounds, based on the expressions of Lemma 5 above, for the expectation and variance. Finally, we bound the expectation over $\boldsymbol{\rho}$ of the total variation distance between the observed probability distribution $\mathcal{H}(\boldsymbol{\rho}, \mathbf{b})$ and the uniform distribution. These bounds will proceed by distinguishing a subset $\Lambda \subset \widehat{G}$ of "bad" representations σ with the undesirable property that the normalized character $|\chi_\sigma(M)/d_\sigma|$ is large; in all cases of interest, these representations will have low dimension.

For a given Λ, we define

$$\lambda = \lambda(M) \triangleq \max_{\sigma\notin\Lambda}\left|\frac{\chi_\sigma(M)}{d_\sigma}\right| \ .$$

We remark that associated with a set Λ and a conjugacy class M of involutions one may immediately compute an upper bound on the ℓ_1-distance between $\mathcal{H}^{\otimes k}(\cdot)$ and $\mathcal{P}^{\otimes k}(\cdot)$. The triangle inequality and Equation 2.1 imply

$$\left\|\mathcal{H}^{\otimes k} - \mathcal{P}^{\otimes k}\right\|_1 \leq k\,\|\mathcal{H} - \mathcal{P}\|_1 \leq 2k\big(\lambda + \mathcal{P}(\Lambda)\big) \ . \qquad (4.7)$$

As we show in Section 5, in the case relevant to Graph Isomorphism this distance is $n^{-O(n)}$. This allows us to assume throughout that the ρ_i are chosen according to the Plancherel measure $\mathcal{P}$ rather than to $\mathcal{H}$, or equivalently, that $\boldsymbol{\rho}$ is chosen according to the Plancherel measure $\mathcal{P}^{\otimes k}$.

4.2 Controlling the expectation

In this section we show that the expected probability distribution $\mathrm{Exp}_m \mathcal{H}^{\otimes m}(\boldsymbol{\rho}, \cdot)$ is close to uniform. First, as we will be concerned with how representations $\boldsymbol{\rho}$ of G^k decompose into irreducible G-representations, we note that for any $\sigma \in \widehat{G}$ and any $I \neq \emptyset$, the expected dimension of the isotypic space corresponding to σ in $\rho_I \otimes \mathbb{1}$, namely d_σ times the multiplicity $a_\sigma^{\rho_I \otimes \mathbb{1}}$, is given by

$$\mathrm{Exp}_{\boldsymbol{\rho}} \frac{a_\sigma^{\rho_I \otimes \mathbb{1}} d_\sigma}{d_{\boldsymbol{\rho}}} = \frac{d_\sigma^2}{|G|} = \mathcal{P}(\sigma) \ , \tag{4.8}$$

if $\boldsymbol{\rho}$ is chosen according to the Plancherel measure [29]. This allows us to show the following bound on the expectation of the involution projector.

Lemma 6. *Let $\Lambda \subset \widehat{G}$, let $\boldsymbol{\rho} = \otimes_{i=1}^k \rho_i$ be chosen according to the Plancherel distribution on $\widehat{G^k}$, let B be an arbitrary basis for $\boldsymbol{\rho}$, and let m be chosen uniformly from a conjugacy class M of involutions. Let $\lambda = \lambda(M)$ be defined as above. Then*

$$\mathrm{Exp}_{\boldsymbol{\rho}} \mathrm{Exp}_{\mathbf{b} \in B} \left| \mathrm{Exp}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2 - \frac{1}{2^k} \right| \leq \lambda + \mathcal{P}(\Lambda) \ .$$

Proof. For any $\boldsymbol{\rho}$ and $\mathbf{b}$, Lemma 5 and the triangle inequality imply that

$$\left| \mathrm{Exp}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2 - \frac{1}{2^k} \right| \leq \frac{1}{2^k} \sum_{I \neq \emptyset} \sum_{\sigma \in \widehat{G}} \left| \frac{\chi_\sigma(M)}{d_\sigma} \right| \left\| \Pi_\sigma^I \mathbf{b} \right\|^2 \ .$$

Pessimistically assuming that $|\chi_\sigma(M)/d_\sigma| = 1$ for all $\sigma \in \Lambda$ and applying the trivial bound $\sum_{\sigma \notin \Lambda} \left\| \Pi_\sigma^I \mathbf{b} \right\|^2 \leq 1$ we conclude that

$$\left| \mathrm{Exp}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2 - \frac{1}{2^k} \right| \leq \lambda + \frac{1}{2^k} \sum_{I \neq \emptyset} \sum_{\sigma \in \Lambda} \left\| \Pi_\sigma^I \mathbf{b} \right\|^2$$

Now observe that for any basis B_ρ of $\boldsymbol{\rho}$ we have

$$\mathrm{Exp}_{\mathbf{b} \in B} \left\| \Pi_\sigma^I \mathbf{b} \right\|^2 = \frac{a_\sigma^{\rho_I \otimes \mathbb{1}} d_\sigma}{d_{\boldsymbol{\rho}}}$$

since $a_\sigma^{\rho_I \otimes \mathbb{1}} d_\sigma$ is the total dimension of the isotypic subspace of $\rho_I \otimes \mathbb{1}$ corresponding to σ. Applying (4.8) completes the proof. □

Corollary 7. *Let Λ and λ be defined as above and let $\boldsymbol{\rho}$ be selected according to the Plancherel distribution. Let $\mathcal{A}(\boldsymbol{\rho}, \mathbf{b}) = \mathrm{Exp}_m \mathcal{H}^{\otimes k}(\boldsymbol{\rho}, \mathbf{b})$ and let $\mathcal{U}$ denote the uniform distribution on B. Then*

$$\mathrm{Exp}_{\boldsymbol{\rho}} \left\| \mathcal{U} - \mathcal{A}(\boldsymbol{\rho}, \cdot) \right\|_1 \leq 2 \cdot 2^k (\lambda + \mathcal{P}(\Lambda)) \ .$$

Proof. Define $\mathcal{I}(\boldsymbol{\rho}, \mathbf{b}) = 2^k \operatorname{Exp}_m \left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2$; note that unless $\mathbf{rk}\ \Pi_m^{\otimes k} = d_{\boldsymbol{\rho}}/2^k$, this is not generally a probability distribution. Then Lemma 6 above asserts that $\operatorname{Exp}_{\boldsymbol{\rho}} \|\mathcal{U} - \mathcal{I}(\boldsymbol{\rho}, \cdot)\|_1 \leq 2^k(\lambda + \mathcal{P}(\Lambda))$. Let $E = \{\boldsymbol{\rho} \in \widehat{G^k} \mid \forall i : \rho_i \notin \Lambda\}$ and notice that as $\boldsymbol{\rho}$ is selected according to the Plancherel distribution, $\Pr[\boldsymbol{\rho} \in E] \geq 1 - k\mathcal{P}(\Lambda)$. When $\boldsymbol{\rho} \in E$, Equation (3.3) implies

$$\mathbf{rk}\ \Pi_m^{\otimes k} = \frac{d_{\boldsymbol{\rho}}}{2^k} \prod_i \left(1 + \frac{\chi_{\rho_i}}{d_{\rho_i}}\right) \in \frac{d_{\boldsymbol{\rho}}}{2^k}\left[(1-\lambda)^k, (1+\lambda)^k\right]$$

and hence $(1-\lambda)^k \mathcal{H}(\boldsymbol{\rho}, \mathbf{b}) \leq \mathcal{I}(\boldsymbol{\rho}, \mathbf{b}) \leq (1+\lambda)^k \mathcal{H}(\boldsymbol{\rho}, \mathbf{b})$. Evidently $\|\mathcal{I}(\boldsymbol{\rho}, \cdot) - \mathcal{H}(\boldsymbol{\rho}, \cdot)\|_1 \leq (1+\lambda)^k - 1 \leq 2^k\lambda$. Pessimistically assuming that this distance is one when $\boldsymbol{\rho} \notin E$ and using the triangle inequality completes the proof. □

4.3 Controlling the variance

We focus now on bounding the projectors contributing to the E^{I_1,I_2} and hence to the variance in Lemma 5 (cf. Equation (4.3)). First, we provide a general bound on the expectation of $|\langle \mathbf{b}, g\mathbf{b}\rangle|^2$ where g ranges over the entire group.

Claim 8. *Let ρ be a representation of a group G acting on a space V and let $\mathbf{b} \in V$. Let g be an element of G chosen uniformly at random. Then*

$$\operatorname{Exp}_g |\langle \mathbf{b}, g\mathbf{b}\rangle|^2 \leq \sum_{\sigma \in \widehat{G}} \frac{\|\mathfrak{I}_\sigma^\rho \mathbf{b}\|^4}{d_\sigma} \ .$$

Proof. Let $\rho \cong \oplus_j \sigma_j$, these σ_j being irreducible, and let $V \cong \oplus V_j$ be the corresponding orthogonal decomposition of V. Write $\mathbf{b} = \sum_j \mathbf{b}_j$ where $\mathbf{b}_j \in V_j$, and $\mathbf{b}_\sigma = \mathfrak{I}_\sigma^\rho \mathbf{b} = \sum_{j:\sigma_j \cong \sigma} \mathbf{b}_j$. This gives

$$\begin{aligned}
\frac{1}{|G|}\sum_{g\in G} |\langle \mathbf{b}, g\mathbf{b}\rangle|^2 &\leq \frac{1}{|G|}\sum_{g\in G}\Big|\sum_j \langle \mathbf{b}_j, g\mathbf{b}_j\rangle\Big|^2 = \frac{1}{|G|}\sum_{g\in G}\sum_{j,k} \langle \mathbf{b}_j, g\mathbf{b}_j\rangle \langle \mathbf{b}_k, g\mathbf{b}_k\rangle^* \\
&= \sum_{j,k} \langle \mathbf{b}_j| \left(\frac{1}{|G|}\sum_{g\in G} |g\mathbf{b}_j\rangle \langle g\mathbf{b}_k|\right) |\mathbf{b}_k\rangle = \sum_\sigma \frac{1}{d_\sigma} \sum_{j,k:\sigma_j,\sigma_k \cong \sigma} |\langle \mathbf{b}_j, \mathbf{b}_k\rangle|^2 \qquad (4.9) \\
&\leq \sum_\sigma \frac{1}{d_\sigma} \sum_{j,k:\sigma_j,\sigma_k\cong\sigma} \|\mathbf{b}_j\|^2 \|\mathbf{b}_k\|^2 = \sum_\sigma \frac{1}{d_\sigma} \|\mathbf{b}_\sigma\|^4 \ , \qquad (4.10)
\end{aligned}$$

as desired. Here we use Schur's lemma [10] in (4.9) and the Cauchy-Schwartz inequality in (4.10). Note that in the inner product of (4.9) we regard $\mathbf{b}_j$ and $\mathbf{b}_k$ as lying in the same copy of σ. □

Lemma 9. *With $\Pi_\sigma^{I_1,I_2}$ defined as in Section 4.1, we have*

$$\sum_{I_1,I_2} \left\|\Pi_\sigma^{I_1,I_2}(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2 \leq 2^k d_\sigma^2 \left(\sum_{I\neq\emptyset} \sum_{\tau\in\widehat{G}} \frac{\left\|\mathfrak{I}_\tau^{\rho_I \otimes \mathbb{1}}\mathbf{b}\right\|^2}{d_\tau}\right) \ .$$

Proof. We can write $\Pi_\sigma^{I_1,I_2}$ as $\mathfrak{I}_\sigma^{(\rho_{I_1}\otimes\mathbb{1})\otimes(\rho_{I_2}\otimes\mathbb{1})^*}$, where $\mathbb{1}$ and $\mathbb{1}^*$ act on $\rho_{\overline{I}_1}$ and $\rho^*_{\overline{I}_2}$ respectively. Using the same notation as in Section 4.1, let g^{I_1,I_2} act on $\mathbf{b}\otimes\mathbf{b}^*$ by multiplying the ith register of $\mathbf{b}$ by m whenever $i\in I_1$, multiplying the ith register of $\mathbf{b}^*$ whenever $i\in I_2$, and leaving the other registers of $\mathbf{b}$ and $\mathbf{b}^*$ unchanged. From the defining expression of Definition 1 we have

$$\left\|\Pi_\sigma^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 = \frac{d_\sigma}{|G|}\sum_{g\in G}\chi_\sigma(g)^*\left\langle \mathbf{b}\otimes\mathbf{b}^*, g^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\right\rangle = \frac{d_\sigma}{|G|}\sum_{g\in G}\chi_\sigma(g)^*\left\langle \mathbf{b}, g^{I_1}\mathbf{b}\right\rangle\left\langle \mathbf{b}, g^{I_2}\mathbf{b}\right\rangle^* \quad .$$

Observe, however, that

$$\sum_{I_1,I_2}\left\|\Pi_\sigma^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 = \frac{d_\sigma}{|G|}\sum_{g\in G}\chi_\sigma(g)^*\sum_{I_1,I_2}\left\langle \mathbf{b}, g^{I_1}\mathbf{b}\right\rangle\left\langle \mathbf{b}, g^{I_2}\mathbf{b}\right\rangle^* = \frac{d_\sigma}{|G|}\sum_{g\in G}\chi_\sigma(g)^*\left|\sum_I\left\langle \mathbf{b}, g^I\mathbf{b}\right\rangle\right|^2$$
$$\leq \frac{d_\sigma^2}{|G|}\sum_{g\in G}\left|\sum_I\left\langle \mathbf{b}, g^I\mathbf{b}\right\rangle\right|^2 \leq 2^k\frac{d_\sigma^2}{|G|}\sum_{g\in G}\sum_I\left|\left\langle \mathbf{b}, g^I\mathbf{b}\right\rangle\right|^2 = 2^k d_\sigma^2\sum_I \mathrm{Exp}_g\left|\left\langle \mathbf{b}, g^I\mathbf{b}\right\rangle\right|^2$$

by the triangle inequality and Cauchy-Schwarz. Finally, we apply Claim 8 to the expectations above and use the fact that as $\|\mathbf{b}\|=1$, $\|\Pi\mathbf{b}\|^4\leq\|\Pi\mathbf{b}\|^2$ for any projection operator Π. □

Then the following lemma bounds the variance of $\left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2$ just as Lemma 6 bounds the expectation.

Lemma 10. *Let $\Lambda\subset\widehat{G}$, let $\boldsymbol{\rho}=\otimes_{i=1}^k\rho_i$ where the ρ_i are independently chosen according to the Plancherel distribution, let B be an arbitrary basis for $\boldsymbol{\rho}$, and let m be chosen uniformly from a conjugacy class M of involutions. Let $\lambda=\lambda(M)$ be defined as above. Then*

$$\mathrm{Exp}_{\boldsymbol{\rho}}\mathrm{Exp}_{\mathbf{b}\in B}\mathrm{Var}_m\left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2 \leq \Delta \triangleq \lambda + \mathcal{P}(\Lambda)\left(\sum_{\tau\in\widehat{G}} d_\tau\right) \quad .$$

Proof. Applying Lemma 9 to control the terms in $E^{I_1,I_2}(\mathbf{b})$ where $\sigma\in\Lambda$, pessimistically assuming that $|\chi_\sigma(M)/d_\sigma|=1$ for all $\sigma\in\Lambda$, and using the obvious bound $\sum_{\sigma\notin\Lambda}\left\|\Pi_\sigma^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2\leq 1$ for the others, we see from (4.5) that

$$\mathrm{Var}_m\left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2 \leq \frac{1}{4^k}\sum_{I_1,I_2\neq\emptyset}E^{I_1,I_2}(\mathbf{b}) \leq \frac{1}{4^k}\sum_{I_1,I_2}\sum_{\sigma\in\widehat{G}}\left|\frac{\chi_\sigma(M)}{d_\sigma}\right|\left\|\Pi_\sigma^{I_1,I_2}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2$$
$$\leq \lambda + \frac{1}{4^k}\sum_{I_1,I_2}\sum_{\sigma\in\Lambda}\left\|\Pi_\sigma^{I_1,I_2}\mathbf{b}\otimes\mathbf{b}^*\right\|^2 = \lambda + \frac{1}{2^k}\left(\sum_{\sigma\in\Lambda}d_\sigma^2\right)\sum_{I\neq\emptyset}\sum_{\tau\in\widehat{G}}\frac{\left\|\mathfrak{I}_\tau^{\rho_I\otimes\mathbb{1}}\mathbf{b}\right\|^2}{d_\tau} \quad .$$

Now we take the expectation of this over the basis B. Since $\mathrm{Exp}_{\mathbf{b}\in B}\left\|\mathfrak{I}_\tau^{\rho_I\otimes\mathbb{1}}\mathbf{b}\right\|^2 = a_\tau^{\rho_I\otimes\mathbb{1}}d_\tau/d_{\boldsymbol{\rho}}$, we have

$$\mathrm{Exp}_{\mathbf{b}\in B}\mathrm{Var}_m\left\|\Pi_m^{\otimes k}\mathbf{b}\right\|^2 \leq \lambda + \frac{1}{2^k}\left(\sum_{\sigma\in\Lambda}d_\sigma^2\right)\sum_{I\neq\emptyset}\sum_{\tau\in\widehat{G}}\frac{a_\tau^{\rho_I\otimes\mathbb{1}}}{d_{\boldsymbol{\rho}}}$$

and Equation (4.8) completes the proof. □

4.4 Bounding the total variation distance

Finally, the next lemma relates the bound of Lemma 10 to the expected variation distance of the observed distribution from the uniform distribution.

Lemma 11. *Let Λ and λ be defined as above, let $\boldsymbol{\rho}$ be selected according to the Plancherel distribution, and let m be uniformly random in its conjugacy class. Let B be a basis for $\boldsymbol{\rho}$ and let $\mathcal{U}$ denote the uniform distribution on B. Then*

$$\operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_m \left\| \mathcal{H}(\boldsymbol{\rho}, \cdot) - \mathcal{U} \right\|_1 \leq 2^k \left[(1-\lambda)^{-k} \sqrt{\Delta} + 3 \cdot (\lambda + \mathcal{P}(\Lambda)) \right]$$

where Δ is defined as in Lemma 10.

Proof. As in Corollary 7, let $\mathcal{A}(\boldsymbol{\rho}, \mathbf{b})$ denote $\operatorname{Exp}_m \mathcal{H}(\boldsymbol{\rho}, \mathbf{b})$. Then we have, analogous to Lemma 7,

$$\begin{aligned}
\operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_m \left\| \mathcal{H}(\boldsymbol{\rho}, \cdot) - \mathcal{A}(\boldsymbol{\rho}, \cdot) \right\|_1 &= \operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_m \sum_{\mathbf{b} \in B} \left| \mathcal{H}(\boldsymbol{\rho}, \mathbf{b}) - \mathcal{A}(\boldsymbol{\rho}, \mathbf{b}) \right| \\
&\leq \operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_m \sqrt{d_{\boldsymbol{\rho}}^2 \operatorname{Exp}_{\mathbf{b} \in B} \left| \mathcal{H}(\boldsymbol{\rho}, \mathbf{b}) - \mathcal{A}(\boldsymbol{\rho}, \mathbf{b}) \right|^2} \leq \operatorname{Exp}_{\boldsymbol{\rho}} \sqrt{\operatorname{Exp}_m d_{\boldsymbol{\rho}}^2 \operatorname{Exp}_{\mathbf{b} \in B} \left| \mathcal{H}(\boldsymbol{\rho}, \mathbf{b}) - \mathcal{A}(\boldsymbol{\rho}, \mathbf{b}) \right|^2} \\
&= \operatorname{Exp}_{\boldsymbol{\rho}} \sqrt{\operatorname{Exp}_{\mathbf{b} \in B} d_{\boldsymbol{\rho}}^2 \operatorname{Var}_m \mathcal{H}(\boldsymbol{\rho}, \mathbf{b})} \leq 2^k (1-\lambda)^{-k} \operatorname{Exp}_{\boldsymbol{\rho}} \sqrt{\operatorname{Exp}_{\mathbf{b} \in B} \operatorname{Var}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2} + k \mathcal{P}(\Lambda) \\
&\leq 2^k (1-\lambda)^k \sqrt{\operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_{\mathbf{b} \in B} \operatorname{Var}_m \left\| \Pi_m^{\otimes k} \mathbf{b} \right\|^2} + k \mathcal{P}(\Lambda) \ .
\end{aligned}$$

The proof is completed by Lemma 10, Corollary 7, and the triangle inequality. □

5 The total variation distance

Having established the generic bounds of the previous sections, it remains simply to apply them to a given group, using a description of its irreducible representations and a choice of the subset Λ. The standard reduction from Graph Isomorphism yields permutations of $2n$ objects, namely the vertices of two graphs of n vertices each. However, rather than all of S_{2n}, it suffices to consider the subgroup $K = S_n \wr \mathbb{Z}_2 \subset S_{2n}$ consisting of permutations which either fix the two vertex sets or swap them.

The irreducible representations of K and their characters are discussed in the Appendix. Our choice of "bad" representations $\Lambda \subset \widehat{K}$ consists of those induced up from representations $\rho \otimes \rho$ of $S_n \times S_n$ with the property that $d_\rho < n^{n/5}$. Simple counting arguments then show that $\lambda \leq n^{-n/5}$, $\mathcal{P}(\Lambda) = n^{-6n/5} e^{O(n)}$, and $\Delta = n^{-n/5} e^{O(n)}$ where Δ is as defined in Lemma 10. With the understanding that $k = n^{O(1)}$, we have $(1-\lambda)^{-k} = 1 + o(1)$ and we find that the expected variation distance in Lemma 11 is

$$\operatorname{Exp}_{\boldsymbol{\rho}} \operatorname{Exp}_m \left\| \mathcal{H}(\boldsymbol{\rho}, \cdot) - \mathcal{U} \right\|_1 \leq 2^k n^{-n/10} e^{O(n)} \ .$$

Thus if $k < Cn \log_2 n$ where C is bounded below $1/10$, this is $n^{-\Omega(n)}$, and by Markov's inequality the probability in $\boldsymbol{\rho}$ and m that $\left\| \mathcal{H}(\boldsymbol{\rho}, \cdot) - \mathcal{U} \right\|_1 > n^{-\Omega(n)}$ is no more than $n^{-\Omega(n)}$. Finally, since by Equation 4.7 $\left\| \mathcal{H}(\cdot) - \mathcal{P}(\cdot) \right\|_1 \leq 2(\lambda + \mathcal{P}(\Lambda)) = n^{-\Omega(n)}$, any event that holds with probability Q in $\mathcal{P}(\cdot)$ holds with probability $Q - n^{-\Omega(n)}$ in $\mathcal{H}(\cdot)$. This completes the proof of Theorem 1; we have made no effort to optimize the constant C.

We remark that these bounds can be established if S_n is replaced with any group G for which a sufficient fraction of the Plancherel measure lies on high-dimensional representations. For any such group, the hidden subgroup problem on $G \wr \mathbb{Z}_2$ requires entangled measurements on $\Theta(n \log n)$ coset states.

Acknowledgments.

This work was supported by the NSF under grants EIA-0218443, EIA-0218563, CCR-0220070, CCR-0220264, and CCF-0524613, and the ARO under grant W911NF-04-R-0009. We are grateful to Sean Hallgren for informing us of his work with Martin Rötteler and Pranab Sen. We thank Tracy Conrad and Sally Milius for their support and tolerance. C.M. also thanks Rosemary Moore for providing a larger perspective.

References

[1] David Bacon, Andrew Childs, and Wim van Dam. Optimal measurements for the dihedral hidden subgroup problem. Preprint, quant-ph/0501044 (2005).

[2] David Bacon, Andrew Childs, and Wim van Dam. From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups. *Proc. 46th Symposium on Foundations of Computer Science*, 2005.

[3] Robert Beals. Quantum computation of Fourier transforms over symmetric groups. *Proc. 29th Annual ACM Symposium on the Theory of Computing*, pages 48–53, 1997.

[4] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory (preliminary abstract). *Proc. 25th Annual ACM Symposium on the Theory of Computing*, pages 11–20, 1993.

[5] Andrew Childs and Paweł Wojcan, On the quantum hardness of solving isomorphism problems as nonabelian hidden shift problems. Preprint, quant-ph/0510185 (2005).

[6] Wim van Dam, Sean Hallgren, and Lawrence Ip. Quantum algorithms for some hidden shift problems. *Proc. 14th ACM-SIAM Symposium on Discrete Algorithms*, pages 489–498, 2003.

[7] Mark Ettinger and Peter Høyer. On quantum algorithms for noncommutative hidden subgroups. Preprint, quant-ph/9807029 (1998).

[8] Mark Ettinger and Peter Høyer and Emmanuel Knill. The quantum query complexity of the hidden subgroup problem is polynomial. *Information Processing Letters*, to appear.

[9] Katalin Friedl, Gábor Ivanyos, Frédéric Magniez, Miklos Santha, and Pranab Sen. Hidden translation and orbit coset in quantum computing. *Proc. 35th ACM Symposium on Theory of Computing*, 2003.

[10] William Fulton and Joe Harris. *Representation Theory: A First Course*. Number 129 in Graduate Texts in Mathematics. Springer-Verlag, 1991.

[11] Michelangelo Grigni, Leonard J. Schulman, Monica Vazirani, and Umesh Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Proc. 33rd ACM Symposium on Theory of Computing*, pages 68–74, 2001.

[12] Lisa Hales and Sean Hallgren. Quantum Fourier sampling simplified. *Proc. 31st Annual ACM Symposium on Theory of Computing*, 1999.

[13] Lisa Hales and Sean Hallgren. An improved quantum Fourier transform algorithm and applications. *Proc. 41st Annual Symposium on Foundations of Computer Science*, 2000.

[14] Sean Hallgren, Alexander Russell, and Amnon Ta-Shma. Normal subgroup reconstruction and quantum computation using group representations. *Proc. 32nd ACM Symposium on Theory of Computing*, pages 627–635, 2000.

[15] Peter Høyer. Efficient quantum transforms. Preprint, quant-ph/9702028 (1997).

[16] Yoshifumi Inui and François Le Gall. An efficient algorithm for the hidden subgroup problem over a class of semi-direct product groups. Proc. EQIS 2004.

[17] Lawrence Ip. Shor's algorithm is optimal. Preprint, 2004.

[18] Gábor Ivanyos, Frédéric Magniez, and Miklos Santha. Efficient quantum algorithms for some instances of the non-abelian hidden subgroup problem. *Int. J. Found. Comput. Sci.* 14(5): 723–740, 2003.

[19] Richard Jozsa. Quantum factoring, discrete logarithms and the hidden subgroup problem. Preprint, quant-ph/0012084 (2000).

[20] Julia Kempe and Aner Shalev. The hidden subgroup problem and permutation group theory. Preprint, quant-ph/0406046 (2004).

[21] S. V. Kerov. *Asymptotic representation theory of the symmetric group and its applications in analysis*. Translated by N. V. Tsilevich. Volume 219 in Translations of Mathematical Monographs. American Mathematical Society, 2003.

[22] Pascal Koiran, Vincent Nesme, and Natacha Portier. A quantum lower bound for the query complexity of Simon's problem. *Proc. of the 32nd International Colloquium on Automata, Languages and Programming*, 2005.

[23] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. Preprint, quant-ph/0302112 (2003).

[24] Cristopher Moore, Daniel Rockmore, and Alexander Russell. Generic quantum Fourier transforms. *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 778–787, 2004.

[25] Cristopher Moore, Daniel Rockmore, Alexander Russell, and Leonard Schulman. The value of basis selection in Fourier sampling: hidden subgroup problems for affine groups. *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1113–1122, 2004.

[26] Cristopher Moore and Alexander Russell. Explicit multiregister measurements for hidden subgroup problems; or, Fourier sampling strikes back. Preprint, quant-ph/0504067 (2005).

[27] Cristopher Moore and Alexander Russell. For distinguishing conjugate hidden subgroups, the pretty good measurement is as good as it gets. Preprint, quant-ph/0501177 (2005).

[28] Cristopher Moore and Alexander Russell and Leonard Schulman. The symmetric group defies Fourier sampling. *Proc. 46th Symposium on Foundations of Computer Science*, pages 479–488 (2005).

[29] Cristopher Moore and Alexander Russell. The symmetric group defies strong Fourier sampling: part II. Preprint, quant-ph/0501066 (2005).

[30] Cristopher Moore and Alexander Russell. Quantum Measurements for Graph Isomorphism Require Entanglement: Tight Results on Multiregister Fourier Sampling Preprint, quant-ph/0510233 (2005).

[31] Jaikumar Radhakrishnan, Martin Rötteler, and Pranab Sen. On the Power of Random Bases in Fourier Sampling: Hidden Subgroup Problem in the Heisenberg Groups. *Proc. 32nd International Colloquium on Automata, Languages and Programming* (2005).

[32] Oded Regev. Quantum computation and lattice problems. *Proc. 43rd Symposium on Foundations of Computer Science*, pages 520–530, 2002.

[33] Martin Rötteler and Thomas Beth. Polynomial-time solution to the hidden subgroup problem for a class of non-abelian groups. Preprint, quant-ph/9812070 (1998).

[34] Jean-Pierre Serre. *Linear Representations of Finite Groups*. Number 42 in Graduate Texts in Mathematics. Springer-Verlag, 1977.

[35] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[36] Daniel R. Simon. On the power of quantum computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997.

[37] A. M. Vershik and S. V. Kerov. Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group. Funk. Anal. i Prolizhen, 19(1):25–36, 1985; English translation, Funct. Anal. Appl., 19:21–31, 1989.

A The group generated by structured involutions

In this section we review the representation theory of the symmetric group S_n, and describe the representations of the subgroup of S_{2n} relevant to Graph Isomorphism. First, recall that the irreducible representations ρ of S_n are labeled by Young diagrams, or equivalently integer partitions $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_t$ such that $\sum_i \lambda_i = n$. The number of irreducible representations is then the partition number $p(n) = e^{O(\sqrt{n})}$.

In the standard reduction from Graph Isomorphism, we consider subgroups $\{1, m\}$ where m is an involution consisting of n disjoint transpositions, matching each vertex in one graph with the corresponding vertex in the other. However, rather than considering all such conjugates in S_{2n}, it makes sense to focus on those involutions m which map $\{1, \ldots, n\}$ to $\{n+1, \ldots, 2n\}$, which we identify with the vertex sets V_1 and V_2 of the two graphs. Such m lie inside a subgroup of S_{2n}: namely, if s denotes a canonical involution $(1\ n+1)(2\ n+2)\ldots(n\ 2n)$, then $m = \alpha^{-1} s \alpha$ where α permutes V_1.

The set of all such involutions generates a subgroup K of S_{2n}. Let $S_{n,n}$ denote the *Young subgroup* $S_{n,n}$ which fixes the sets V_1 and V_2; then K is the subgroup generated by $S_{n,n}$ and s. Algebraically, K is the *wreath product* $S_n \wr \mathbb{Z}_2$, and can also be written as a semidirect product $K = (S_n \times S_n) \rtimes \mathbb{Z}_2$. If $\alpha, \beta \in S_n$ and $t \in \mathbb{Z}_2$, we denote by $((\alpha, \beta), t)$ the element which applies α to V_1 and β to V_2, and then applies s^t. Note that $|K| = 2n!^2 = n^{2n} e^{-O(n)}$.

We can determine K's irreducible representations and their characters as follows. For two irreducible representations ρ and σ of S_n, let $\rho \boxtimes \sigma$ denote their tensor product as a representation of $S_{n,n} \cong S_n \times S_n$. We consider the induced representation $\tau_{\{\rho,\sigma\}} = \mathrm{Ind}_{S_{n,n}}^{K}(\rho \boxtimes \sigma)$ and denote its character $\chi_{\{\rho,\sigma\}}$. It is easy to see that

$$\chi_{\{\rho,\sigma\}}\big(((\alpha,\beta),t)\big) = \begin{cases} 0 & \text{if } t = 1 \\ \chi_\rho(\alpha)\chi_\sigma(\beta) + \chi_\sigma(\alpha)\chi_\rho(\beta) & \text{if } t = 0 \ ; \end{cases}$$

as the notation suggests, this depends only on the multiset $\{\rho, \sigma\}$. An easy computation shows that $\langle \chi_{\{\rho,\sigma\}}, \chi_{\{\rho,\sigma\}} \rangle = 1 + \delta_{\rho,\sigma}$. Thus, if $\rho \ncong \sigma$, then $\tau_{\{\rho,\sigma\}}$ is irreducible of dimension $2d_\rho d_\sigma$; while if $\rho \cong \sigma$ then it decomposes into two irreducible representations of dimension d_ρ^2,

$$\tau_{\{\rho,\rho\}} \cong \tau_{\{\rho,\rho\},\mathbb{1}} \oplus \tau_{\{\rho,\rho\},\pi}$$

where $\mathbb{1}$ and π are the trivial and sign representations, respectively, of $\mathbb{Z}_2$. Each of these irreducible representations acts on $V_\rho \otimes V_\rho$, the vector space supporting the action of $\rho \boxtimes \rho$. Both of them realize the element $((\alpha, \beta), 0)$ as the linear map $\rho(\alpha) \otimes \rho(\beta)$, while $\tau_{\{\rho,\rho\},\mathbb{1}}$ and $\tau_{\{\rho,\rho\},\pi}$ realize the element $((1,1),1)$ as the maps which send $\mathbf{u} \otimes \mathbf{v}$ to $\mathbf{v} \otimes \mathbf{u}$ and $-\mathbf{v} \otimes \mathbf{u}$ respectively. The characters of these representations are

$$\chi_{\{\rho,\rho\},\mathbb{1}}((\alpha,\beta),t) = \begin{cases} \chi_\rho(\alpha) + \chi_\rho(\beta) & \text{if } t = 0 \\ \chi_\rho(\alpha\beta) & \text{if } t = 1 \end{cases}, \qquad \chi_{\{\rho,\rho\},\pi}((\alpha,\beta),t) = \begin{cases} \chi_\rho(\alpha) + \chi_\rho(\beta) & \text{if } t = 0 \\ -\chi_\rho(\alpha\beta) & \text{if } t = 1 \end{cases}.$$

In particular, since m is of the form $((\alpha, \alpha^{-1}), 1)$, we have the normalized characters

$$\frac{\chi_{\{\rho,\rho\},\mathbb{1}}(m)}{d_{\{\rho,\rho\},\mathbb{1}}} = \frac{1}{d_\rho}, \quad \frac{\chi_{\{\rho,\rho\},\pi}(m)}{d_{\{\rho,\rho\},\pi}} = -\frac{1}{d_\rho} \tag{A.1}$$

and $\chi_{\{\rho,\sigma\}}(m) = 0$ for all $\rho \ncong \sigma$.

We remark that this construction of the irreducible representations and their characters works for any group of the form $G \wr \mathbb{Z}_2$. In particular, the normalized characters of the involutions that "swap" the two copies of G are either 0 or $\pm 1/d_\rho$ for some $\rho \in \widehat{G}$.

If we choose Λ to consist of those $\tau_{\{\rho,\rho\},\mathbb{1}}$ and $\tau_{\{\rho,\rho\},\pi}$ such that $d_\rho < n^{n/5}$, then by by (A.1) we have $\lambda \leq n^{-n/5}$. Since there are at most $p(n)^2$ irreducible representations of K we have

$$\mathcal{P}(\Lambda) = \sum_{\tau \in \Lambda} d_\tau^2 / |K| \leq p(n)^2 \, n^{4n/5} / |K| = n^{-6n/5} e^{O(n)} .$$

Similarly, since no irreducible representations of K can have dimension greater than $\sqrt{|K|}$, the quantity Δ defined in Lemma 10 is bound by

$$\Delta = \lambda + \frac{\mathcal{P}(\Lambda)}{|K|} \left(\sum_{\tau \in \widehat{G}} d_\tau \right) \leq n^{-n/5} + n^{-6n/5} \, p(n)^4 \sqrt{|K|} \, e^{O(n)} = n^{-n/5} e^{O(n)} .$$